

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2003, 04:41:21 ; Search time 0.674157 seconds  
(Without alignments)

553.709 Million cell updates/sec

Title: US-09-698-781-17

Perfect score: 44

Sequence: 1 TLPVLLFL 9

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt.40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	44	100.0	245 1 CRS3_HUMAN	P54108 homo sapien
2	35	79.5	587 1 NDC2_RAT	P70545 rattus norv
3	34	77.3	247 1 DB83_HUMAN	P57088 homo sapien
4	34	77.3	247 1 DB83_RAT	O92142 rattus norv
5	34	77.3	301 1 REIS_TODPA	P23820 todardes p
6	34	77.3	741 1 NMSC_FLARA	O32238 flavaria ra
7	33	75.0	245 1 CRS3_HORSE	O19010 equus cabal
8	33	75.0	261 1 GILT_HUMAN	P13284 homo sapien
9	33	75.0	365 1 RECF_CHLMR	O96Kw5 chlamydia m
10	33	75.0	121 1 YDGF_ECOLI	O84077 chlamydia t
11	32	72.7	405 1 CYB_RHORB	P23134 rhodospirill
12	32	72.7	409 1 PEXC_PICPA	O01961 pichia past
13	32	72.7	519 1 PPBT_CHICK	O92058 gallus gall
14	32	72.7	571 1 CVCA_PEA	P13915 pisum sativ
15	32	72.7	100 1 YJPO_YEAST	P47004 saccharomyc
16	31	70.5	154 1 YBGA_MERYA	P81329 methanococc
17	31	70.5	170 1 Y019_BORBU	O51051 botryella bu
18	31	70.5	197 1 I17C_HUMAN	O490M4 homo sapien
19	31	70.5	221 1 API1_SOLTU	O94448 solanum tub
20	31	70.5	221 1 API7_SOLTU	O41448 solanum tub
21	31	70.5	247 1 NUGM_WHEAT	O02500 triticum ae
22	31	70.5	324 1 NULM_PISCO	P24995 plasteer oc
23	31	70.5	428 1 KINB_BACSU	O08430 bacillus su
24	31	70.5	433 1 PSNI_XENLA	O19576 xenopus lae
25	31	70.5	456 1 PTSE_KLEPN	P27219 klebsiella
26	31	70.5	456 1 PTSE_KLEPN	P084768 salmoneila
27	31	70.5	456 1 PSNI_SALTU	P49768 homo sapien
28	31	70.5	467 1 PSNI_MICMU	P73802 microcebus
29	31	70.5	467 1 PSNI_MOUSE	P49769 mus musculu
30	31	70.5	467 1 PSNI_MOUSE	P49769 mus musculu
31	31	70.5	468 1 PSNI_BOVIN	P97887 rattus norv
32	31	70.5	478 1 PSNI_BOVIN	O92197 bos taurus
33	31	70.5	499 1 MEP2_YEAST	P41948 saccharomyc

34	31	70.5	517 1 HSF2_MOUSE	P38533 mus musculu
35	31	70.5	536 1 HSF2_HUMAN	O03933 homo sapien
36	31	70.5	543 1 P69_MYCGE	P47533 mycoplasma
37	31	70.5	564 1 HSF2_CHICK	P38530 gallus galli
38	31	70.5	679 1 PAN3_YEAST	P36102 saccharomyc
39	31	70.5	784 1 TLR2_YEAST	O991f8 cricetus
40	31	70.5	784 1 TLR2_MOUSE	O991f8 mus musculu
41	31	70.5	1204 1 DMB1_VZVD	P09246 varicella-z
42	31	70.5	1816 1 LMA4_HUMAN	O16363 homo sapien
43	30	68.2	145 1 YNR9_YEAST	P53880 saccharomyc
44	30	68.2	153 1 MAL_CANFA	O28296 canis famil
45	30	68.2	181 1 DSBF_PASMU	O9cpm6 pasteurella

## ALIGNMENTS

```

RESULT 1
ID CRS3_HUMAN STANDARD: PRT: 245 AA.
AC P54108; Q15512;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteine-rich secretory protein-3 precursor (CRISP-3) (SGP28 protein).
GN CRISP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE-96270732; PubMed-8665901;
RA Kretzschmar J., Haendler B., Eberspacher U., Roostermann D.,
RA Donner P., Schlemming W.-D.;
RT "The human cysteine-rich secretory protein (CRISP) family. Primary
RT structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3."
RL Eur. J. Biochem. 236:827-836(1996).
RN [2]
RP MEDLINE-96186934; PubMed-8601434;
RA Kjeldsen L., Cowland J.B., Johnson A.H., Borregaard N.;
RT "SGP28, a novel matrix glycoprotein in specific granules of human
RT neutrophils with similarity to a human testis-specific gene product
RT and a rodent sperm-coating glycoprotein."
RL FEBS Lett. 380:246-250(1996)
CC -I- SUBCELLULAR LOCATION: SECRETED; IN NEUTROPHILS, LOCALIZED IN
CC -I- SPECIFIC GRANULES.
CC -I- TISSUE SPECIFICITY: SALIVARY GLAND, PANCREAS AND PROSTATE >
CC EPIDIDYMIS, OVARY, THYMUS AND COLON.
CC -I- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MAMMALIAN SCP/PPX1;
CC INSECTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sib-sdb.ch).
CC
CC EMBL: X95240; CAA64527.1;
CC EMBL: X94373; CAA63984.1;
CC HSSP: P04284; ICRB.
CC InterPro: IPR001283; Allrgn_V5/Tpx1.
CC Pfam: PF00188; SCP_1.
CC PRINTS: PR00837; V5TPXLIKE.
CC ProDom: PD000542; Allrgn_V5/Tpx1; 1.
CC SMART: SM00198; SCP_1.
CC PROSITE: PS01009; SCP_AG5_PRL_SCP_1; 1.
CC PROSITE: PS01010; SCP_AG5_PRL_SCP_2; 1.
CC Glycoprotein: Signal; Multigene family; Polymorphism.
CC SIGNAL 1 20 POTENTIAL.

```

```

FT CHAIN 21 245 CYSTEINE-RICH SECRETORY PROTEIN-3.
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 106 106 S->P (IN DBSNP:495335).
FT VARIANT 134 134 /FTID-VAR.011718.
FT VARIANT 134 134 A->S (IN DBSNP:1864312).
FT SEQUENCE 245 AA; 27630 MW; BADD79CB/AE9E5F9 CRC64;

Query Match 100.0%; Score 44; DB 1; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TLFPVLFL 9
Db 2 TLFPVLFL 10

RESULT 2
NDC2_RAT STANDARD; PRT: 587 AA.
ID NDC2_RAT
AC P70545;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intracellular sodium/dicarboxylate cotransporter (Na(+)/dicarboxylate
cotransporter).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mistral;
RA Khatir I.A., Kovacs S.V.B., Forstner J.F.;
RA MetLife I.A., Kovacs S.V.B., Forstner J.F.;
RT Cloning of the cDNA for a rat intestinal Na+/dicarboxylate
cotransporter reveals partial sequence homology with a rat intestinal
mucin.
RL Biochem. Biophys. Acta 1309:58-62(1996).
CC -1- FUNCTION: COTRANSPORT OF SODIUM AND DICARBOXYLATES SUCH AS
SUCCINATE AND CITRATE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE NDC/P/PHO87 FAMILY OF TRANSPORTERS.
CC NDC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U51153; AAB97095.1;
DR InterPro: IPR001898; Na_sulph_symp.
DR Pfam: PF00939; Na_sulph_symp; 1.
DR PROSITE: PS01271; NA_SULFATE; 1.
KW Transport; Transmembrane; Sodium transport; Symport.
FT TRANSMEM 13 33 POTENTIAL.
FT TRANSMEM 53 73 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 326 346 POTENTIAL.
FT TRANSMEM 367 387 POTENTIAL.
FT TRANSMEM 407 427 POTENTIAL.
FT TRANSMEM 445 465 POTENTIAL.
FT TRANSMEM 477 497 POTENTIAL.
FT TRANSMEM 506 526 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
SO SEQUENCE 587 AA; 63531 MW; F9B74F921BDC8712 CRC64;

Query Match 79.5%; Score 35; DB 1; Length 587;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 2 LFPVLFL 9
Db 61 LFPVLFL 68

RESULT 3
DB83_HUMAN STANDARD; PRT: 247 AA.
ID DB83_HUMAN
AC P57088;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DB83 protein.
DE DB83.
GN DB83.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shitatori A., Sudo H.,
RA Wagatsuma M., Hosokawa T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Minomiyu K., Iwayanagi T.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0121 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK001387; BAA91665.1;
DR InterPro: IPR005344; UPF0121.
DR Pfam: PF03661; UPF0121; 1.
KW Transmembrane.
FT TRANSMEM 32 52 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
SO SEQUENCE 247 AA; 27951 MW; 9C481A13ECCACB16 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 247;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TLFPVLFL 8
Db 111 SIFVLFL 118

RESULT 4
DB83_RAT STANDARD; PRT: 247 AA.
ID DB83_RAT
AC Q92142;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DB83 protein.
DE DB83.
GN DB83.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE-99087491; PubMed-9672456;
RA Nakadai T., Kishimoto T., Kokura K., Ohkawa N., Makino Y.,
RA Muramatsu M., Tamura T.;
RT "Cloning of a novel rat gene, DB83, that encodes a putative membrane
RT protein."
RL DNA Res. 5:315-317(1998).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE LIVER AND
CC SIGNIFICANTLY IN BRAIN, LUNGS AND KIDNEYS.
CC -1- SIMILARITY: BELONGS TO THE UPF0121 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB006135; BAA75068.1; -
DR InterPro; IPR005344; UPF0121.
DR Pfam; PF03661; UPF0121; 1.
KW Transmembrane.
FT TRANSMEM 32 52 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.
SQ SEQUENCE 247 AA; 27983 MW; 0D07BE5911CAD6AE CRC64;

Query Match 77.3%; Score 34; DB 1; Length 247;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LFPVLFLF 8
Db 111 SLFPLVLF 118

RESULT 5
REIS_TODPA STANDARD; PRT; 301 AA.
AC P23820;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
DE Retinochrome (Retinal photoisomerase).
OS Todarodes pacificus (Japanese flying squid).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoida; Teuthoidea;
OC Oegopsida; Ommastrephidae; Todarodes.
OX NCBI_TaxID=6637;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Retina;
RA MEDLINE-91032043; PubMed-2226795;
RA Hara-Nishimura I., Matsumoto T., Mori H., Nishimura M., Hara R.,
RA Hara T.;
RT "Cloning and nucleotide sequence of cDNA for retinochrome, retinal
RT photoisomerase from the squid retina."
RL FEBS Lett. 271:106-110(1990).
RN [2]
RN SEQUENCE OF 1-15 AND 114-128.
RC TISSUE-Retina;
RA Uematsu J., Hara-Nishimura I., Wada K., Matsubara H., Hara T.;
RT "Amino-terminal sequence of squid retinochrome."
RL Photobiochem. Photobiophys. 13:197-201(1986).
RN [3]
RN SEQUENCE OF 274-281, AND RETINAL BINDING SITE.
RP TISSUE-Retina;
RA MEDLINE-94063090; PubMed-8243675;
RA Hara-Nishimura I., Kondo M., Nishimura M., Hara R., Hara T.;
RT "Amino acid sequence surrounding the retinal-binding site in
RT retinochrome of the squid, Todarodes pacificus."
RL FEBS Lett. 335:94-98(1993).
CC -1- FUNCTION: RETINOCROME IS CAPABLE OF ACTING AS AN EFFECTIVE

```

```

CC CATALYST IN THE LIGHT TO CONVERT VARIOUS ISOMERS OF RETINAL
CC INTO 11-CIS, THE FORM THAT IS REQUIRED BY OPSIN TO RESYNTHESIZE
CC RHODOPSIN.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MYELOID MEMBRANE.
CC -1- TISSUE SPECIFICITY: MAINLY STORED IN MYELOID BODIES OF THE INNER
CC SEGMENTS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC OPSIN SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X57143; CAA40422.1; -
DR PIR; S12864; S12864.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001760; Opsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL1; FALSE_NEG.
DR PROSITE; PS50262; G-PROTEIN_RECEP_FL2; 1.
KW PROSITE; PS00238; OPSIN; 1.
KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;
KW G-protein coupled receptor.
FT DOMAIN 1 17 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 18 43 1 (POTENTIAL).
FT DOMAIN 44 54 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 55 76 2 (POTENTIAL).
FT DOMAIN 77 94 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 95 120 3 (POTENTIAL).
FT DOMAIN 121 132 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 133 153 4 (POTENTIAL).
FT DOMAIN 154 180 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 181 208 5 (POTENTIAL).
FT DOMAIN 209 230 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 231 255 6 (POTENTIAL).
FT DOMAIN 256 264 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 265 289 7 (POTENTIAL).
FT DOMAIN 290 301 CYTOPLASMIC (POTENTIAL).
FT BINDING 275 275 RETINAL CHROMOPHORE.
FT CARBOHYD 170 170 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 301 AA; 33490 MW; CBB37317486B27FC CRC64;

Query Match 77.3%; Score 34; DB 1; Length 301;
Best Local Similarity 75.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LFPVLFLF 9
Db 280 LFPVLFLF 287

RESULT 6
NU5C_FLARA STANDARD; PRT; 741 AA.
AC Q32238;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE NADH-plastoquinone oxidoreductase chain 5, chloroplast (EC 1.6.5.3).
GN NDH.
OS Flaveria ramosissima.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eunsterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Flaveria.
OX NCBI_TaxID=41578;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE-96036088; PubMed-7479788;
RA Kim K.J., Jansen R.K.;
RT "ndhf sequence evolution and the major clades in the sunflower
family.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10379-10383(1995).
CC -1- CATALYTIC ACTIVITY: NADH + plastoquinone - NAD(+) + plastoquinol.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L39465; AAC37748.1;
DR InterPro: IPR003916; NADHoxrxd5.
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR001516; Oxidored_q1_N.
DR Pfam: PF00361; Oxidored_q1; 1.
DR Pfam: PF00662; Oxidored_q1_N; 1.
DR Pfam: PF01010; Oxidored_q1_C; 1.
DR PRINTS: PRO1434; NADHoxrxd5.
KW Oxidoreductase; NAD; Plastoquinone; Chloroplast.
SO SEQUENCE 741 AA; 83948 MW; E0BD194ABEF0B9FB CRC64;

Query Match 77.3%; Score 34; DB 1; Length 741;
Best Local Similarity 62.5%; Pred. No. 46;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 LFPVLLFL 9
Db 549 LFPVLLFL 556

RESULT 7
CRS3_HORSE STANDARD; PRT; 245 AA.
ID CRS3_HORSE
AC O19010;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cysteine-rich secretory protein-3 precursor (CRISP-3).
GN CRISP3.
OS Equus caballus (Horse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
CC NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amputula;
RA MEDLINE-9842318; PubMed-9748582;
RA Schampony A., Gentzel M., Wolfes H., Ralda M., Neumann U.,
RA Toepfer-Petersen E.;
RT Toepfer CRISP-3: primary structure and expression in the male genital
RT tract.
RL Biochim. Biophys. Acta 1387:206-216(1998).
CC -1- SUBCELLULAR LOCATION: SECRETED; IN NEUTROPHILS, LOCALIZED IN
CC SPECIFIC GRANULES (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE SALIVARY GLAND, IN THE
CC AMPULLA AND THE SEMINAL VESICLE.
CC -1- INVERTS AG3/AG5; FUNGI SC7/SC14 AND PLANTS PR-1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL; AJ001400; CA004729.1;
DR HSSP; P04284; ICEE.
DR InterPro: IPR001283; Allrgn_V5/TPx1.
DR Pfam; PF00188; SCP; 1.
DR PRINTS; PR00837; V5TPX1IKE.
DR PRODOM; PD000542; Allrgn_V5/TPx1; 1.
DR SMART; SM00198; SCP; 1.
DR PROSITE; PS01009; SCP_AGS_PRL_SC7_1; 1.
DR PROSITE; PS01010; SCP_AGS_PRL_SC7_2; 1.
KW Signal: Multigene family.
FT SIGNAL 1 22
FT CHAIN 23 245
SO SEQUENCE 245 AA; 27308 MW; 893A4EB7F402BA22 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 245;
Best Local Similarity 87.5%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LFPVLLFL 9
Db 3 LFPVLLFL 10

RESULT 8
GILT_HUMAN STANDARD; PRT; 261 AA.
ID GILT_HUMAN
AC P13284; Q9UL08;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gamma-interferon inducible lysosomal thiol reductase precursor
DE (Gamma-interferon-inducible protein IP-30).
GN IFI30 OR GILT.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-8829888; PubMed-3136170;
RA Luster A.D., Weinshank R.L., Felman R., Ravetch J.V.;
RT "Molecular and biochemical characterization of a novel gamma-
RT interferon-inducible protein."
RL J. Biol. Chem. 263:12036-12043(1988).
RN [2]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RX MEDLINE-20105543; PubMed-10639150;
RA Arunachalam B., Phan U.T., Geuze H.J., Cresswell P.;
RT "Enzymatic reduction of disulfide bonds in lysosomes: characterization
RT of a gamma-interferon-inducible lysosomal thiol reductase (GILT).";
RL Proc. Natl. Acad. Sci. U.S.A. 97:745-750(2000).
CC -1- FUNCTION: CLEAVES DISULFIDE BONDS IN PROTEINS BY REDUCTION. MAY
CC FACILITATE THE COMPLETE UNFOLDING OF PROTEINS DESTINED FOR
CC LYSOSOMAL DEGRADATION. MAY BE INVOLVED IN MHC CLASS II-RESTRICTED
CC ANTIGEN PROCESSING.
CC -1- SUBCELLULAR LOCATION: Lysosomal.
CC -1- INDUCTION: EXPRESSED CONSTITUTIVELY IN ANTIGEN-PRESENTING CELLS
CC AND INDUCED BY IFN-GAMMA IN OTHER CELL TYPES.
CC -1- PTM: N-GLYCOSYLATED. SUGAR CHAINS CONTAIN MANNOSE-6-PHOSPHATE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03909; AAA36105.1;
DR EMBL; AF097362; AA04618.1;
DR PIR; A43708; A43708.
DR Gene; HGNC:5398; IFI30.
MM; 604664;

```

```

DR InterPro: IPR001238; Recf.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02463; SMC_N; 1.
DR TIGRfams: TIGR00611; recf; 1.
DR PROSITE: PS00617; Recf_1; 1.
DR PROSITE: PS00618; Recf_2; 1.
KW DNA damage; DNA replication; DNA-binding; SOS response; DNA repair;
KW ATP-binding; Complete proteome.
FT NP_BIND 30
SQ SEQUENCE 365 AA: 41321 MW: F097CE97810C8958 CRC64:
-----
Query Match
Best Local Similarity 75.0%; Score 33; DB 1; Length 365;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 2 LEPUVF 8
DB 110 LEPUVF 116
-----
RESULT 10
REFC_CHLTR STANDARD: PRT: 365 AA.
AC 084077;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA replication and repair protein recf.
DE RECF OR CT074.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UV-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koehn E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis."
RL Science 282:754-759(1998).
CC -1- FUNCTION: THE RECF PROTEIN IS INVOLVED IN DNA METABOLISM; IT IS
CC REQUIRED FOR DNA REPLICATION AND NORMAL SOS INDUCIBILITY. RECF
CC BINDS PREFERENTIALLY TO SINGLE-STRANDED, LINEAR DNA. IT ALSO SEEMS
CC TO BIND ATP (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE RECF FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: AE001282; AAC67665.1;
CC InterPro: IPR001238; Recf.
CC InterPro: IPR003395; SMC_N.
CC Pfam: PF02463; SMC_N; 1.
CC TIGRfams: TIGR00611; recf; 1.
CC PROSITE: PS00617; Recf_1; 1.
CC PROSITE: PS00618; Recf_2; 1.
KW DNA damage; DNA replication; DNA-binding; SOS response; DNA repair;
KW ATP-binding; Complete proteome.
FT NP_BIND 30
SQ SEQUENCE 365 AA: 41246 MW: CC058A9475638A0B CRC64:
-----
Query Match
Best Local Similarity 75.0%; Score 33; DB 1; Length 365;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0

```

OY	2	LFPVLF 8
D6	110	LFPVLF 116

```

RESULT 11
YDGF_ECOLI STANDARD: PRT; 121 AA.
ID YDGF_ECOLI STANDARD: PRT; 121 AA.
AC P77412:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ydGF.
GN YDGF OR B1600 OR Z2594 OR ECS2306.
OS Escherichia coli, and
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OX Escherichia.
XX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alta H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubraman S.,
RA Yamamoto Y., Takeuchi K., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Heckelt J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Djmalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blatner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[4]
RN RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
RA Ikeda T., Takami H., Honda T., Sasaki K., Ogawara N., Yasunaga T.,
RA Kunihira S., Shibata H., Hattori M., Shinagawa H.,
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genetic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[5]
RN RP SUBCELLULAR LOCATION: Integral membrane protein (Potential).
[6]
RN RP SIMILARITY BELONGS TO THE SMALL MOLETCULOG RESISTANCE (SMR)
PROTEIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AE000255; AAC74672.1; -
CC DR EMBL; D90803; BAA15339.1; -
CC DR EMBL; D90802; BAA15334.1; -
CC DR EMBL; AE005383; AAC656587.1; -
CC DR EMBL; AF002558; BAB35729.1; -
CC DR Ecogen; EG1327; ydgE.
CC DR InterPro; IPR000390; DUF7.
CC KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
CC FT TRANSMEM 2 22 POTENTIAL.
CC FT TRANSMEM 32 52 POTENTIAL.
CC FT TRANSMEM 55 75 POTENTIAL.
CC FT TRANSMEM 82 102 POTENTIAL.
CC SQ SEQUENCE 121 AA; 13115 MW; B24330AF5BBI81B5 CRC64;
CC -----
Query Match 72.7%; Score 32; DB 1; Length 121;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 73 TLFSYLLF 80
QY 1 TLFPYLLF 8
1111111
CYB_RHORU STANDARD; PRT; 405 AA.
ID CYB_RHORU P23134;
AC P23134;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 30-May-2000 (Rel. 39, Last annotation update)
DE Cytochrome B.
DE PEBB.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Rhodospirillum.
CX NCBI_Taxid=1085;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.
RC STRAIN=FR1;
RX MEDLINE=91094774; PubMed=2176269;
RA Majewski C., Trebst A.;
RT "The pet genes of Rhodospirillum rubrum: cloning and sequencing of
RT the genes for the cytochrome bcl-complex.";
RL Mol. Gen. Genet. 224:373-382(1990).
CC -1- FUNCTION: COMPONENT OF THE UBINOULO-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -1- COPACITOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X55387; CAA39059.1; -
CC PIR; S12257; CBOFR.
CC InterPro; IPR000179; Cyt_b_b6.
CC Pfam; PF00032; cytochrome_b_c1.
CC Pfam; PF00033; cytochrome_b_n_1.

```

DR PROSITE: PS00192; CYTOCHROME\_B\_HEME; 1.  
 DR PROSITE: PS00193; CYTOCHROME\_B\_OO; 1.  
 KW Electron transport; Respiratory chain; Heme; Transmembrane.  
 FT METAL 94 94 IRON 1 (HEME B562 AXIAL LIGAND).  
 FT METAL 108 108 IRON 2 (HEME B566 AXIAL LIGAND).  
 FT METAL 195 195 IRON 2 (HEME B566 AXIAL LIGAND).  
 FT METAL 209 209 IRON 1 (HEME B566 AXIAL LIGAND).  
 SQ SEQUENCE 405 AA; 46369 MW; F19452EDA18BE6F CRC64;

Query Match 72.7%; Score 32; DB 1; Length 405;  
 Best Local Similarity 85.7%; Pred. No. 63;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 LFPVLLF 8  
 DB 197 LFPVLLF 203

RESULT 13  
 PEPC\_PICPA STANDARD; PRT; 409 AA.  
 AC 001961;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Peroxisome assembly protein PAS10 (Peroxin-12).  
 GN PEX12 OR PAS10.  
 OS Pichia pastoris (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetaceae; Pichia.  
 OX NCBI\_TaxID=4922;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96283626; PubMed=8670828;  
 RA Kallish J.E., Gould S.J., Morrell J.C., Mihalik S.J., Smith B.,  
 RA Clegg J.M., Gould S.J.;  
 RT "Characterization of a novel component of the peroxisomal protein  
 RT import apparatus using fluorescent peroxisomal proteins".  
 RL EMBO J. 15:3275-3285(1996).  
 CC -1- FUNCTION: REQUIRED FOR PROTEIN IMPORT INTO PEROXISOMES.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Peroxisomal.  
 CC -1- SIMILARITY: BELONGS TO THE PEROXIN 12 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U58140; AAC49402.1;  
 DR InterPro: IPR001841; Znf\_Lfing.  
 DR SMART: SM00184; RING; 1.  
 KW Transmembrane; Peroxisome.  
 FT TRANSMEM 171 191  
 FT TRANSMEM 267 287 POTENTIAL.  
 SQ SEQUENCE 409 AA; 47591 MW; 1F2DB79A957C3B7F CRC64;

Query Match 72.7%; Score 32; DB 1; Length 409;  
 Best Local Similarity 66.7%; Pred. No. 64;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TLFPVLLF 9  
 DB 273 TLFPASIFL 281

RESULT 14  
 PPBT\_CHICK STANDARD; PRT; 519 AA.  
 ID PPBT\_CHICK  
 AC 092058;  
 DT 15-JUL-1999 (Rel. 38, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Alkaline phosphatase, tissue-nonspecific isozyme precursor  
 DE (EC 3.1.3.1) (AP-TNAP) (Liver/bone/kidney isozyme).  
 GN ALPL.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96107571; PubMed=8563025;  
 RA Crawford K., Weisig H., Binette F., Millan J.L., Goetlink P.F.;  
 RT "Tissue-nonspecific alkaline phosphatase participates in the  
 RT establishment and growth of feather germs in embryonic chick skin  
 RT cultures".  
 RL Dev. Dyn. 204:48-56(1995).  
 RL -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an  
 RL alcohol + phosphate.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U91908; AAA92562.1;  
 DR HSSP: P00634; 2ANH.  
 DR InterPro: IPR001952; Alk-phosphatse.  
 DR Pfam: PF00245; alk-phosphatase; 1.  
 DR PRINTS: PR00113; ALKPHPTASE.  
 DR ProDom: PD001668; Alk-phosphatse; 1.  
 DR SMART: SM00098; alkPc; 1.  
 DR PROSITE: PS00123; ALKALINE\_PHOSPHATASE; 1.  
 KW Hydrolyase; zinc; Magnesium; Phosphorylation; Transmembrane;  
 KW Multigene family; Glycoprotein; GPI-anchor; Signal.  
 FT SIGNAL 1 16  
 FT CHAIN 17 519 ALKALINE PHOSPHATASE, TISSUE-NONSPECIFIC  
 FT MOD\_RES 109 109  
 FT ACT\_SITE 109 109 PHOSPHORYLATION (BY SIMILARITY).  
 FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 229 229 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 278 278 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 429 429 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 519 AA; 56760 MW; 7934C0ECC3B17889 CRC64;

Query Match 72.7%; Score 32; DB 1; Length 519;  
 Best Local Similarity 77.8%; Pred. No. 80;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TLFPVLLF 9  
 DB 507 TLFPVLLF 515

RESULT 15  
 CVCA\_PEA STANDARD; PRT; 571 AA.  
 ID CVCA\_PEA  
 AC P13915;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Convicillin precursor.  
 GN CVCA.

```

OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Feltham First;
RX MEDLINE=88326208; PubMed=3415641;
RA Bown D., Ellis T.H.N., Gatehouse J.A.;
RT "The sequence of a gene encoding convicillin from pea (Pisum sativum
RT L.) shows that convicillin differs from vicillin by an insertion near
RL Biochem. J. 251:717-726(1988).
CC -!- FUNCTION: SEED STORAGE PROTEIN.
CC -!- SUBCELLULAR LOCATION: CYTOLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN
CC BODIES.
CC -!- SIMILARITY: BELONGS TO THE 7S SEED STORAGE PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X06398; CAA29695.1; -.
DR PIR: S00566; S00566.
DR HSSP: P02853; 2PHL.
DR InterPro: IPR001113; Seedstore_7s.
DR Pfam: PF00546; Seedstore_7s; 1.
DR Pfam: PF02808; Seedstore_7s_C; 1.
KW Seed storage protein; Multigene family; Signal.
FT SIGNAL 1 28
FT CHAIN 29 571 CONVICILIN.
SQ SEQUENCE 571 AA; 66989 MW; 749CFEBB2D16D57B CRC64;

```

Search completed: March 14, 2003, 05:40:56  
 Job time : 1.67416 secs

```

OY 3 FVLLFL 9
   11:1111
DB 9 FVLLFL 15

```

Query Match 72.7%; Score 32; DB 1; Length 571;  
 Best Local Similarity 85.7%; Pred. NO. 88;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;